

REMARKS

Claims 1-18 are pending in the application. Claims 1, 4-8 and 11-16 are rejected under 35 U.S.C. § 102(b) as being anticipated by Kannes, U.S. Patent No. 5,382,972 ("Kannes"). Claims 2-3,¹ 9-10 and 17-18 are rejected under 35 U.S.C. § 103(a) as being unpatentable over Kannes in view of Iizawa, U.S. Patent No. 5,801,756 ("Iizawa"). Applicant respectfully traverses the rejections by the Examiner on the basis that Kannes, Iizawa, and the other prior art of reference fail to teach or suggest certain novel features of the claims of the present invention, as discussed more fully below.

1. Summary of the Invention

In a communications system that supports conference calls that include an audio portion and a video portion, the present invention provides a method for selecting a primary video image from a plurality of video images. A novel feature of this method is determining an amount of audio data generated by each participant on the conference call. Preferably, the amount of data is determined by counting a number of audio packets or by counting an amount of audio samples in audio packets. After the amount of audio data is determined, a dominant audio participant is selected, if possible, based upon the amount of audio data from each participant. Preferably, a dominant audio participant is selected if the difference in the amount of audio data is substantial, the dominant audio participant being the participant generating the most audio data. If a dominant

¹Applicant's attorney had a telephone conversation with the Examiner on September 16, 2003, regarding the Official Action dated July 17, 2003. In that conversation, Applicant's attorney raised the issue that Claim 2-3 were not mentioned in the detailed action and therefore the reasons for rejecting those claims was not understood. The Examiner indicated this was the result of a typographical error and that the reference to "Claims 4-5" as it appears on page 5, paragraph 4, line 1 should have been "Claims 2-3". This response is made as though the typographical error is rectified as specified by the Examiner.

audio participant is selected, a primary video image associated with that dominant audio participant is selected as the primary video image for display.

2. The Claims Are Patentably Distinguishable Over Kannes.

Claims 1, 4-8 and 11-16 are rejected under 35 U.S.C. § 102(b) as being anticipated by Kannes. Applicant asserts that claim 1 is allowable over Kannes because it recites,

In a communications system that supports conference calls that include an audio portion and a video portion, a method for selecting a primary video image from a plurality of video images, the method comprising . . . determining an *amount* of audio data generated by each participant of a plurality of participants in a conference call; selecting a dominating audio participant . . . based upon the *amount* of audio data generated by each participant . . . and selecting a primary video image based on the dominating audio participant (emphasis added).

Kannes does not disclose “determining an amount of audio data generated by each participant of a plurality of participants in a conference call . . . and selecting a dominating audio participant . . . based upon the amount of audio data generated by each participant . . .”

For selecting a primary video image, Kannes discloses “a voice sensitive switching means” which “selects a video signal” (col. 5, lines 59-61). However, the voice sensitive switching means in Kannes does not determine an amount of audio data in order to select a video signal. Instead, “the voice sensitive switching means . . . select[s] for display . . . the video image associated with an active one of the local . . . microphones” (col. 5, lines 65-67). More specifically, the microphone signals are connected to an active microphone signal generation unit (col. 13, lines 37-39). This unit monitors the microphone input signals “in a conventional manner to determine which is active at any instant” (col. 13, lines 40-42).

As evident in the examples set out in Kannes, the voice sensitive switching means selects merely those microphones which are active, regardless of the amount of voice data coming through

these microphones (col. 6, lines 1-20; col. 16, line 9). For example, if two local conferees speak simultaneously into their microphones, the software used in Kannes will select, as having the active microphone, first one and then the other conferee in a preset alternating pattern (col. 6, lines (col. 6, lines 4-9). Thus, the software in Kannes selects participants, and in turn, video images of the participants, based merely on the fact that a particular speaker's microphone is active, *i.e.*, the participant has to make a sound above a predetermined qualitative threshold for a defined duration of time through the microphone (col. 16, lines 8-11).

In contrast, the present invention, as recited in claim 1, selects the primary video image of a participant based upon the amount of audio data generated by that participant and not based upon the active or passive state of a given microphone. Unlike Kannes, the present invention, as recited in claim 1, compares the amount of audio data generated by each participant, and based upon that comparison, selects a primary video image of the participant.

In conclusion, claim 1 is patentable over Kannes, at least because Kannes does not disclose “determining an amount of audio data generated by each participant of a plurality of participants in a conference call” and “selecting a dominating audio participant . . . based upon the amount of audio data generated by each participant.” Claims 2-5 depend on claim 1 and are therefore allowable for at least the reasons given above with respect to claim 1.

Independent claim 6 is allowable over Kannes for much the same reasons as applicable to claim 1. In addition, claim 6 is allowable because it recites the following additional step, determining whether a difference between an amount of audio data generated by one participant . . . and an amount of audio data generated by other participants . . . exceeds a predetermined threshold.

With respect to the Examiner's reasons for rejecting Claim 6, Kannes discloses a “predetermined minimum amplitude” and a “threshold microphone amplitude” (col. 15, line 56;

col. 16, line 9). These amplitudes refer to what it takes to make a microphone active, *i.e.*, turn it on, as already distinguished with reference to claim 1. However, once a given microphone is active, Kannes does not determine a difference in the amount of audio data generated by participants, as recited in claim 6. Furthermore, Kannes does not select between the participants based upon the amount of audio data generated by each participant.

Kannes relies on timing differences, *i.e.*, who started talking first, in differentiating between participants (col. 6, lines 1-9). For example, in Kannes, if two conferees attempt to speak nearly simultaneously into their microphones, the software will select the microphone of the first person to speak (col. 6, lines 1-4). If two conferees speak simultaneously, the software will alternate the image of both speakers, regardless of the speaker who generates more audio data (col. 6, lines 4-9). Because Kannes does not disclose determining a difference in the amount of audio data generated by the various participants, claim 6 is allowable. Claims 7-12 depend on claim 6 and are therefore allowable for at least the reasons given above with respect to claim 6.

The Examiner rejected claim 13 for reasons similar to the rejections of claims 1 and 6. Claim 13, however, like claims 1 and 6, is patentable over Kannes because it recites an apparatus with a “processor that determines an amount of audio data generated by each participant . . . in a conference call.” As previously discussed with respect to claims 1 and 6, Kannes does not recite a method or an apparatus capable of determining an amount of audio data generated by each participant in a conference call. Hence, claim 13 is allowable. Claims 14-16 depend on claim 13 and are therefore allowable for at least the reasons given above with respect to claim 13.

3. Claims 2-3, 9-10 and 17-18 are patentable over Kannes in view of Iizawa.

Claims 2-3, 9-10 and 17-18 depend on claims 1, 6, and 13, respectively, and are therefore allowable on this basis alone, as discussed above. And, these claims are allowable for the following additional reasons. Each of the claims recites “counting a number of audio packets generated by each participant” or “counting an amount of audio samples in audio packets” or a similar limitation. Applicants submit that the prior art of record, and particularly the Iizawa reference cited by the Examiner, do not teach such a limitation.

With respect to the Examiner’s comments on Iizawa, it is noted that Iizawa uses coders-decoders (“decoders”) to transmit speech and images of participants in a video conference and a controller for processing the speech and image signals from the decoders (col. 2, lines 20-24). The controller includes a count table, which stores utterance times of the participants (col. 2, lines 44-45). The controller measures utterance/ non-utterance times of the participants and selects image signals from the decoders “in accordance with the utterance times of speakers” (col. 4, lines 37-38).

In contrast to the utterance time counting method disclosed in Iizawa, the present invention, as recited in claims 2-3, 9-10 and 17-18, “count[s] a number of audio packets generated by each participant” and “count[s] an amount of audio samples in audio packets.” Neither Kannes nor Iizawa either separately or in combination, discloses or even suggests counting a number of audio packets or counting an amount of audio samples in audio packets. Thus, claims 2-3, 9-10 and 17-18 are allowable.

CONCLUSION

Claims 1-18 are allowable over the cited references. An indication of allowance is solicited at an early date.

Respectfully submitted,



Reginald J. Hill

Reg. No. 39,225

JENNER & BLOCK, LLC
One IBM Plaza
Chicago, Illinois 60611
(312) 222-9350
Date: September 29, 2003

#977526